基于大数据回溯新冠肺炎的扩散趋势及中国对疫情的控制研究

赵序茅[1]　李欣海[2,3]　聂常虹[2*]

1 兰州大学 生态学创新研究院 草地农业生态系统国家重点实验室 兰州 730000
2 中国科学院动物研究所 北京 100101
3 中国科学院大学 北京 100094

摘要　自2019年12月以来，湖北省武汉市出现新型冠状病毒肺炎（简称“新冠肺炎”），随后疫情在全国范围内暴发。相关学者对此次疫情扩散趋势做了大量研究，但基于模型的估算普遍存在高估传染系数和感染人群的问题。基于此，利用大数据回溯新冠肺炎在全国扩散的趋势和传染系数，得出结论：（1）在扩散前期（2020年1月24号之前）武汉输出的人口是全国各城市（武汉除外）新冠肺炎扩散的主要威胁，之后每个省、自治区、直辖市的人口与确诊病例数的关系逐渐增强，体现出本地传播的趋势；（2）利用SEIR模型拟合武汉和北京两地新冠肺炎确诊病例得出真实传播系数（R_0），揭示武汉的R_0高于北京，考虑到人为控制因素，2019年12月1日—2020年2月9日两地的实际的R_0均大幅度下降。基于此，对疫情期间存在的问题进行针对性建议，并从数据上论证了中国政府对于疫情扩散强有力的控制能力，对阻止国内及世界范围的疫情扩散作出积极贡献。

关键词　新冠肺炎，大数据，武汉，北京

DOI　10.16418/j.issn.1000-3045.20200210002

冠状病毒是自然界广泛存在的一种病毒，因在电镜下像王冠而得名，可引起人和动物呼吸道、消化道、神经系统疾病[1]。自2019年12月以来，湖北省武汉市集中发生了多起病毒性肺炎，2020年1月8日查明病原为一种新型冠状病毒，经过基因测序被确定为新型冠状病毒肺炎（简称“新冠肺炎”）；该病毒与SARS冠状病毒具有相似的传播途径，其原始病毒寄生在蝙蝠体内[1,2]。截至2020年2月10

*通讯作者
修改稿收到日期：2020年2月10日；预出版日期：2020年2月17日
① 人民日报-丁香园：全国疫情实时动态（https://ncov.dxy.cn/ncovh5/view/pneumonia_peopleapp?scene=126&clicktime=1579832412&from=timeline&isappinstalled=0）。

日，中国境内确诊 40 235 例，疑似 23 589 例，治愈 3 326 例，死亡 909 例①。疫情发生后，国内外科研人员对新型冠状病毒的起源和感染性[3-5]、进化关系[6]、致病机制[7-9]、宿主[10]、扩散模式、传播动力、传播模型和预测[11-19]、防控建议[20]等做了大量研究和报道。这对疫情的判断、决策、治疗起到积极的作用。与此同时，一些针对疫情扩散的模型，普遍存在低估中国政府对疫情的管控能力，高估计感染人数和传染系数的情况。基于此，本文基于大数据和实时更新的新冠肺炎确诊人数，回溯疫情的扩散趋势，从数据上揭示中国政府对疫情的管控效果，并对疫情期间存在的一些问题提供针对性建议。

1 基于大数据分析新冠病毒扩散趋势

基于百度地图大数据②对武汉输出人口和全国确诊病例进行线性相关分析：湖北省除外，各省份的人口数和从武汉输出的人口数对各省累积确诊病例数的解释度分别为 50% 和 10%，即各城市中50%的确诊病例数与本地人口相关，10%的确诊病例数与武汉输出的人口相关（图1）。新冠肺炎在不同的扩散阶段其扩散趋势不同：在开始阶段，其传播量与 2019 年12月30日—2020 年 1月20日武汉输出人口密切相关，相关系数为 0.71。该相关系数以航空出行的人数为例，根据 2019 年12月30日—2020 年 1月20日从武汉天河国际机场国内航班出行人数③和 2020 年 1月22日全国各城市新冠肺炎确诊病例数[3]进行相关性分析得出的；如果算上从武汉离开的所有交通方式，其相关系数会更高。从数据上论证了新冠肺炎的源头在武汉地区，全国各地城市（除武汉外）71% 的感染病例和从武汉输出的人口相关。2020 年 1月24日，这个相关系数降到 0.56，之后逐渐下降，全国各省、自治区、直辖市的人口与确诊数的关系逐渐增强，确诊病例与城市人口规模呈显著的正相关。与此同时，新冠肺炎呈现本地扩散的趋势，以武汉为中心向湖北境内各个城市快速扩散（图 2）。基于新冠肺炎的扩散模式，自 2020 年 1月23日以来，湖北省的交通管制措施有效

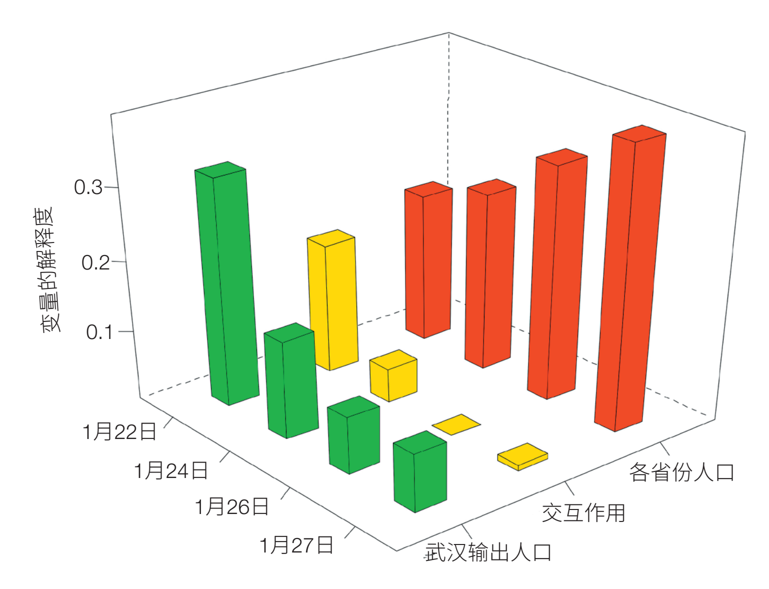


图 1　线性模型统一量化各省份人口、武汉旅客对确诊数的解释度

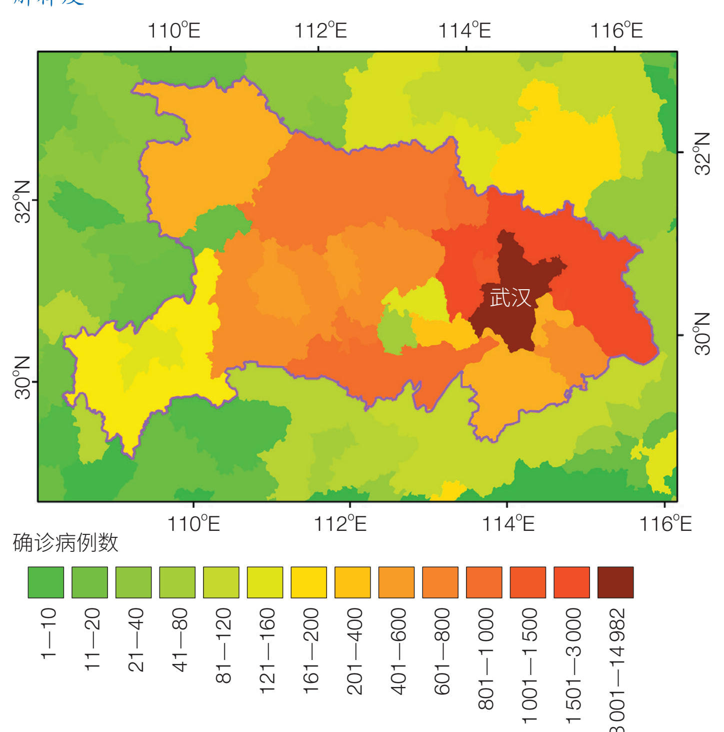


图 2　新冠肺炎在武汉发病的周边扩散趋势（数据更新到 2020 年 2 月 9 日）

② 百度地图慧眼（https://qianxi.baidu.com/?city=420100）。
③ 畅说民航（https://mp.weixin.qq.com/s/7ynWYxB-s7nfz7rmjBLSpQ）。

地遏制了疫情向湖北省外的扩散，缓解了其他城市的压力。

2 中国政府对疫情扩散的控制

传统的传染病模型分析疫情低估了人为因素对于疫情的控制，和现在确诊病例进行对比，之前学者的推算无法接近真实的情形[11-19]。我们利用 SEIR 模型[21]及每天的确诊人数进行拟合，得出真实的传染系数 R_0，之后对北京和武汉的疫情扩散进行对比分析。

SEIR 模型（Hethcote 2000）如下：

$$\frac{dS}{dt}=-\beta\frac{I}{N}S;$$

$$\frac{dE}{dt}=\beta\frac{I}{N}S-\alpha E;$$

$$\frac{dI}{dt}=\alpha E-vI;$$

$$\frac{dR}{dt}=vI。$$

这里 S 代表易感人群，E 代表感染后处在潜伏期的人群，I 代表感染人群，R 代表恢复健康不会再被感染人群，t 代表第一例感染者出现之后的天数，N 代表总人口，β 代表患者 1 天感染的人数，α 代表潜伏期，v 代表恢复率。

结果显示：2019 年 12 月1日到—2020 年1月23日武汉新冠肺炎的传染系数 R_0 达到 5.75，而当 2020 年 1 月 23 日武汉实施交通管制之后，R_0 下降到 2.5；北京的确诊患者中一半直接来源于武汉[22]，根据每日确诊的患者计算，R_0 为 1.5，远低于武汉。

我们的分析佐证了中国各级政府的管控对于疫情的积极影响。从 2020 年 1 月 24 日开始，为了防止疫情迅速蔓延，湖北省至少 18 个城市（县）宣布采取"封城"措施，这从源头上控制了疫情的扩散。据交通运输部消息，2020 年 1 月 10 日—2 月 6 日，全国铁路、道路、水路、民航累计发送旅客同比下降 36.9%。中国具有"集中力量办大事"的制度优势，可以调度全国的力量投入疫情防控。例如，紧急调集专家、医务人员、物资进驻武汉，这对控制新冠肺炎继续向其他省份扩散，对稳定、阻隔疫情的传播意义重大。中国政府在短时间内修建并投入使用雷神山医院、火神山医院、方舱医院，对有效缓解武汉本地的疫情作出了积极努力。

2020 年 1 月 25 日所有发现新冠肺炎病例的省份均已启动重大突发公共卫生事件一级响应④，对疑似患者进行严格排查。基于大数据开发的软件程序可以精确查找确诊、疑似患者所乘坐的车次，以及与确诊或疑似患者的距离，由此可对潜在感染人员进行排查和隔离，对疫情的防控和排查起到关键性作用⑤⑥。为了防止春节后人员流动对于病毒扩散的影响，政府宣布延长春节假期。我国政府强有力的执行能力，最大限度地减少了新冠肺炎的扩散。

3 疫情期间出现的问题和建议

当前，中国政府采取的一系列措施有效地控制住疫情的扩散。目前，72.2% 的新冠肺炎确诊患者在湖北省，且湖北省外每日新增患者呈现下降趋势，这是一个积极的信号。全国防控的重点依旧在湖北。随着火神山医院、雷神山医院、方舱医院的交付使用，以及检测手段的改进，湖北地区的疫情将会得到最大限度的控制和缓解。建议继续加大对湖北地区的支持力度，从源头上控制疫情的扩散，缓解其他省份的压力。除湖北省外，新冠肺炎的确诊数和城市人口数高

④ 人民日报 - 丁香园：全国疫情实时动态（https://ncov.dxy.cn/ncovh5/view/pneumonia_peopleapp?scene=126&clicktime=1579832412&from=timeline&isappinstalled=0）。

⑤ 百度地图慧眼（https://qianxi.baidu.com/?city=420100）。

⑥ 畅说民航（https://mp.weixin.qq.com/s/7ynWYxB-s7nfz7rmjBLSpQ）。

度相关，因此大城市要继续保持严控。此外，新冠肺炎具有长时间的潜伏期，要严密关注春节后的返程人口流动对疫情扩散带来的影响，这直接关乎疫情结束时间。

3.1 当前存在的问题

当前，各个城市对于疫情的防控已经取得积极效果，但是疫情期间依然存在5个方面的问题。

（1）*舆情问题*。舆情是抗击疫情的第二战场。在此次疫情期间，一些学者仅仅依靠模型估算，忽视中国政府对疫情的管控能力，夸大疫情的严重性。例如，根据模型计算中国感染人口在2020年2月4日预计感染例数超过19万（区间为13万—27万），疫情将在武汉之外的上海、北京、广州、重庆、成都等地暴发[23]。国内一些媒体对此进行跟踪报道后，造成一定的舆情恐慌；然而，现实中并未发生。再如，印度学者在bioRxiv上发文怀疑新型冠状病毒里含有HIV病毒的基因片段，被国内自媒体报道后曾经造成一定程度的恐慌。

（2）*信任危机*。包括少数政府官员缺少公信力，以及科研人员公众之间存在信任危机。研究表明：专业知识并不是决定谁最值得信任的决定因素，公众更愿意相信代表自己最大利益的人，这甚至比了解专业知识要重要3—4倍[24]。公众不相信科学家，并不是质疑科学家的专业知识，而是质疑科学家是否与其有着共同的利益。对于突发风险问题的沟通，信任既是一个认知过程，也是一个社会过程。它不仅取决于信息的接收质量，还取决于传播者的特征、社会角色及其与受众的关系[25]。因此，在突发事件上，政府官员和科研人员只有对公众坦诚、负责，站在公众的利益上，才能获得民众的认可与信任。

（3）*心理恐惧问题*。疫情期间，公众普遍产生恐惧心理，不仅对疫情本身恐惧，更担心疫情的传染力、致死率及其后遗症等问题。恐惧心理不仅会对自身健康造成影响，还会影响国家经济。根据行为经济学理论，恐惧会造成对市场消息过度反应，刺激非理性经济行为，严重影响经济市场和金融市场。

（4）*部分地区出现防控过度*。对于疫情需要高度重视，但是也不能违背人性防控过度。诸如，医务人员无法回家、阻断正常的交通道路等问题。防控过度不仅影响正常的社会秩序，还会加重恐慌情绪。

（5）*社会治理及应急机制有待完善*。此次疫情反映出：① 已有的公共安全防护体系在应对突发性事件时缺少有效预警机制，在社会风险控制、公共资源分配、基层社会治理、物资调度、信息真实性及透明度等方面缺少稳定而健全的运转体系；② 国家公职人员在面对突发疫情时，在公共服务、管理水平上存在不足，不能有效提供社会保护，造成社会管理失效，在一定程度上造成政府公信力缺失。

3.2 政策建议

结合对新冠肺炎的扩散趋势和传染系数的分析，以及疫情期间出现的一系列问题，提出5点建议。

（1）*加强舆论引导和监督*。政府相关部门要积极引导舆论，发布真实、可靠的消息，做到信息及时公开，避免被动辟谣。同时，警惕国内外资本夸大疫情、借机炒作中国市场，扰乱中国经济秩序。媒体发布相关疫情信息之前，多询问相关领域的专业人士，确保信息真实、可靠。

（2）*改善信任危机*。疫情期间折射出的信任危机，相关部门要敢于承认、敢于担当、敢于问责。如果一味地回避只会加重信任危机。科研人员要对自己的研究负责，谨慎发表一些可能带来不利于社会团结稳定的信息，疫情期间更应加强监督和同行评议。

（3）*加强公众心理疏导，减少恐惧*。针对公众的恐慌心理应及时进行科普和心理疏导。疫情期间的科普要以解决实际问题为第一要义，不能为了吸引眼球去科普。此外，科普要接地气，要主动出击去解决问题，要以公众通俗易懂的语言去讲述。如果讲清楚新冠病毒的起源、发病机制、致死率等问题，就可以大

幅减少恐慌。同时，要警惕一些打着科普旗号，捕风捉影、制造恐慌的行为。

（4）科学开展防控。针对过度防控问题，首先应对各个城市面临的疫情危机进行正确、科学地评估。本文结论表明，目前各个城市的新冠肺炎的病例数和城市的人口规模高度相关。基于此，城市的疫情防控不能只看新增确诊病例数和累计确诊病例数，还要结合城市人口和面积综合考虑其传染的危险等级。对于一些疫情威胁低的地区，如一些西部城市，当每日增加确诊病例数趋于稳定或停止，当地医疗条件足以应对的情况下，应重点管控人员的跨省流动，适度放松本省内城市间的人员流动，在做好个人防控的基础上及时恢复正常的生产、生活秩序。

（5）完善治理体系。此次疫情是对我国治理体系和能力的一个考验，政府的作为有令人赞叹之处，也有让人痛心之时。应尽早完善应急体系的职责、社会管理机制，以及和政府各部门的联动机制。当前，应尽快将中国应急体系中已经成建制的力量纳入抗击疫情工作中[26]。各级卫生健康委员会系统，包括疾病控制中心、医院及政府其他职能部门应尽快建立和完善稳定健全的协调机制，使应急管理部门的作用充分、有效、合理发挥。

参考文献

1 Xu X, Chen P, Wang J, et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. Science China Life Sciences, 2020, DOI: 10.1007/s11427-020-1637-5.

2 Ji W, Wang W, Zhao X, et al. Homologous recombination within the spike glycoprotein of the newly identified coronavirus may boost cross-species transmission from snake to human. Journal of Medical Virology, 2020, DOI: 10.1002/jmv.25682.

3 Zhou P, Yang X L, Wang X G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature, 2020, DOI: 10.1038/s41586-020-2012-7.

4 Benvenuto D, Giovannetti M, Ciccozzi A, et al. The 2019 - new coronavirus epidemic: Evidence for virus evolution. Journal of Medical Virology, 2020, DOI: 10.1002/jmv.25688.

5 Wu F, Zhao S, Yu B, et al. Complete genome characterisation of a novel coronavirus associated with severe human respiratory disease in Wuhan, China. bioRxiv, 2020, DOI: 10.1101/2020.01.24.919183.

6 Paraskevis D, Kostaki E G, Magiorkinis G, et al. Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. Infection, Genetics and Evolution, 2020, 79: 104212. DOI: 10.1016/j.meegid.2020.104212.

7 Letko M, Munster V. Functional assessment of cell entry and receptor usage for lineage B β-coronaviruses, including 2019-nCoV. bioRxiv, 2020, DOI: 10.1101/2020.01.22.915660.

8 Xu Z, Peng C, Shi Y, et al. Nelfinavir was predicted to be a potential inhibitor of 2019-nCov main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. bioRxiv, 2020, DOI: 10.1101/2020.01.27.921627.

9 Tian X, Li C, Huang A, et al. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. bioRxiv, 2020, DOI: 10.1101/2020.01.28.923011.

10 Guo Q, Li M, Wang C et al. Host and infectivity prediction of Wuhan 2019 novel coronavirus using deep learning algorithm. bioRxiv, 2020, DOI: 10.1101/2020.01.21.914044.

11 Chen T, Rui J, Wang Q, et al. A mathematical model for simulating the transmission of Wuhan novel Coronavirus. bioRxiv, 2020, DOI: 10.1101/2020.01.19.911669.

12 Riou J, Althaus C L. Pattern of early human-to-human transmission of Wuhan 2019-nCoV. bioRxiv, 2020, DOI:

10.1101/2020.01.23.917351.

13 Zhao S, Lin Q, Ran J, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. International Journal of Infectious Diseases, 2020, DOI: 10.1016/j.ijid.2020.01.050.

14 Shen M, Peng Z, Xiao Y, et al. Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. bioRxiv, 2020, DOI: 10.1101/2020.01.23.916726.

15 Chen Z, Zhang W, Lu Y, et al. From SARS-CoV to Wuhan 2019-nCoV outbreak: Similarity of early epidemic and prediction of future trend. bioRxiv, 2020, DOI: 10.1101/2020.01.24.919241.

16 Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine, 2020, DOI: 10.1056/NEJMoa2001017.

17 Liu T, Hu J, Kang M, et al. Transmission dynamics of 2019 novel coronavirus (2019-nCoV). bioRxiv, 2020, DOI: 10.1101/2020.01.25.919787.

18 Zhang C, Wang M. Origin time and epidemic dynamics of the 2019 novel coronavirus. bioRxiv, 2020, DOI: 10.1101/2020.01.25.919688.

19 Ming W, Huang J, Zhang C J P. Breaking down of healthcare system: Mathematical modelling for controlling the novel coronavirus (2019-nCoV) outbreak in Wuhan, China. bioRxiv, 2020, DOI: 10.1101/2020.01.27.922443.

20 Shao P, Shan Y. Beware of asymptomatic transmission: Study on 2019-nCoV prevention and control measures based on extended SEIR model. bioRxiv, 2020, DOI: 10.1101/2020.01.28.923169.

21 Hethcote H W. The mathematics of infectious diseases. Siam Review, 2000, 42: 599-653.

22 Wu J T, Leung K, Leung G M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. The Lancet, 2020, DOI: 10.1016/S0140-6736(20)30260-9.

23 Read J M, Bridgen J R, Cummings D A, et al. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. medRxiv, 2020, DOI: 10.1101/2020.01.23.20018549.

24 Eiser J R, Stafford T, Henneberry J, et al. “Trust me, I’m a Scientist (Not a Developer)”: Perceived expertise and motives as predictors of trust in assessment of risk from contaminated land. Risk Analysis, 2009, 29(2): 288-297.

25 赵斌. 疫情高发，大众为什么还不信任科学家？[2020-02-10]. http://blog.sciencenet.cn/blog-502444-1217726.html.

26 陈安. 对当前肺炎疫情应急的八条建议. [2020-02-10]. http://blog.sciencenet.cn/blog-53483-1216183.html.

Backtracking Transmission of COVID-19 in China Based on Big Data Source, and Effect of Strict Pandemic Control Policy

ZHAO Xumao[1] LI Xinhai[2,3] NIE Changhong[2*]

（1 State Key Laboratory of Grassland Agro-Ecosystem, Institute of Innovation Ecology, Lanzhou University, Lanzhou 730000, China;

2 Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China;

3 University of Chinese Academy of Sciences, Beijing 100049, China）

Abstract In December 2019, COVID-19 appeared and started transmission in local population in Wuhan, Hubei Province. We analyzed the spread process of COVID-19 and found that imported number of passengers from Wuhan before the city closure is the main threat to other cities in China, whereas later on local transmission in those cities gradually become the main force of virus transmission. Based on SEIR model, we found that the basic reproductive number R_0 for Wuhan is much higher than that of Beijing. When pandemic control measures (traffic control, holiday extension, 14-day-long quarantine, etc.) are taken into account, the R_0 dropped substantially. China’s progressive pandemic control policy ensures the situation under control, and the timely situation reporting and data sharing greatly contribute to the whole world fighting against this novel coronavirus.

Keywords COVID-19, big data, Wuhan, Beijing



赵序茅 兰州大学生态学创新研究院青年研究员，中国科学院动物研究所博士，国际灵长类协会会员，中国科普作家协会会员，北京作家协会会员。主要研究领域包括：全球气候变化生态学、保护生物学。在国内外发表学术论文10余篇，出版科普书籍13部，多次获得省部级奖项。E-mail: zhaoxm@lzu.edu.cn

ZHAO Xumao Research Professor, Institute of Innovation Ecology, Lanzhou University. He achieved Ph.D. degree from Institute of Zoology, Chinese Academy of Sciences (CAS), in 2019. He is also the member of International Primate Association, member of China Science Writers Association, and member of Beijing Writers Association. His current research interests include climate change ecology and conservation biology. From 2014 to 2020, he has published more than 10 scientific papers and 13 books about popular science and won provincial and ministerial awards for many times. E-mail: zhaoxm@lzu.edu.cn

* Corresponding author



聂常虹　中国科学院动物研究所党委书记、副所长，中国科学院大学教授、博士生导师，中国科学院大学中国政府和社会资本合作（PPP）研究中心学术委员会副主任，中华人民共和国联合国教科文组织全国委员会咨询专家，全国知识管理标准化技术委员会委员。主要从事科技财政、科技金融、科技成果转化、宏观经济与产业政策等方面的研究，主持财政部、国家发展和改革委员会、科学技术部、国家自然基金委员会等部委重点课题多项。在《新华文摘》《人民日报》《管理评论》《宏观经济研究》《财政研究》和*Journal of Systems Science and Complexity*等国家核心期刊和SCI期刊发表学术论文50余篇，出版合著4部，专著2部；10余篇论文或研究报告获得部级一等奖或二等奖。E-mail: Chnie@ioz.ac.cn

NIE Changhong　Deputy Director of Institute of Zoology (IOZ), Chinese Academy of Sciences (CAS), Researcher and Ph.D. Supervisor at University of Chinese Academy of Sciences (UCAS), Deputy Chair of Academic Committee of Public-Private-Partnership (PPP) Research Center, UCAS, Consultant Expert of National Commission of the People's Republic of China for UNESCO, and member of the National Knowledge Management Standardization Technical Committee. She is mainly engaged in the research of science and technology finance, transformation of scientific and technological achievements, macroeconomics, and industrial policies. She has presided over several projects of the Ministry of Finance, the National Development and Reform Commission, the Ministry of Science and Technology, and the National Natural Science Foundation of China. More than 10 of her papers or official documents have won the first prize or second prize of Ministry level, more than 50 of her academic papers have been published in SCI indexed journals or core journals such as *Xinhua Digest*, *People's Daily*, *Management Review*, *Macroeconomics Research*, *Public Finance Research*, and *Journal of Systems Science and Complexity*. She has also authored 2 monographs and co-authored 4 other monographs.
E-mail: Chnie@ioz.ac.cn

■责任编辑：岳凌生